

Evolutionary Capacitance May Be Favored by Natural Selection

Joanna Masel¹

Department of Biological Sciences, Stanford University, Stanford, California 94305 and Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721

Manuscript received January 4, 2005
Accepted for publication April 6, 2005

ABSTRACT

Evolutionary capacitors phenotypically reveal a stock of cryptic genetic variation in a reversible fashion. The sudden and reversible revelation of a range of variation is fundamentally different from the gradual introduction of variation by mutation. Here I study the invasion dynamics of modifiers of revelation. A modifier with the optimal rate of revelation m_{opt} has a higher probability of invading any other population than of being counterinvaded. m_{opt} varies with the population size N and the rate θ at which environmental change makes revelation adaptive. For small populations less than a minimum cutoff N_{min} , all revelation is selected against. N_{min} is typically quite small and increases only weakly, with $\theta^{-1/2}$. For large populations with $N > 1/\theta$, m_{opt} is $\sim 1/N$. Selection for the optimum is highly effective and increases in effectiveness with larger $N \gg 1/\theta$. For intermediate values of N , m_{opt} is typically a little less than θ and is only weakly favored over less frequent revelation. The model is analogous to a two-locus model for the evolution of a mutator allele. It is a fully stochastic model and so is able to show that selection for revelation can be strong enough to overcome random drift.

ONE of the major puzzles of evolutionary biology is how sufficient phenotypic variation can be generated and maintained to form the basis of an adaptive response to novel environments. A population with very little phenotypic variation will be well adapted to a stable environment in the short term, but may be unable to respond to a change in the environment.

In a well-adapted population almost all newly introduced variation is likely to be deleterious. Variation is eliminated through selection, at the cost of a mutation load. Selective pressure to reduce the load tends to favor the reduction of variation by a variety of means, including reducing the mutation rate (SNIEGOWSKI *et al.* 2000) and buffering against the effects of mutations that do occur (WAGNER 1996; NOWAK *et al.* 1997; ESHEL and MATESSI 1998; RICE 1998, 2002; VAN NIMWEGEN *et al.* 1999; WILKE *et al.* 2001; KRAKAUER and PLOTKIN 2002; DE VISSER *et al.* 2003). This last mechanism is known as canalization and can lead to the buildup of hidden or cryptic genetic variation.

In recent years, a number of mechanisms have been shown to tap into the pool of cryptic genetic variation, revealing heritable phenotypic variation. These include partial loss of function of the heat-shock protein Hsp90 (RUTHERFORD and LINDQUIST 1998; QUEITSCH *et al.* 2002) and the appearance of the yeast prion [*PSI*⁺] (TRUE and LINDQUIST 2000). In each case, genetic variation that was previously hidden is “turned on” at a single

stroke. It remains on for a number of generations, either because of the continuation of environmental stress or because of epigenetic inheritance, but it can be turned off later.

These mechanisms have the potential to promote evolvability, in the sense of revealing potentially adaptive phenotypic variation at a time when it might most be needed. They act as “capacitors.” Following this analogy, variation gradually produced by mutation is stored in a hidden form by the capacitor. When needed, this variation can be released, and the occasional revelation of latent variation may lead to evolutionary innovations.

It is very controversial, however, whether evolutionary capacitors may be the product of natural selection for increased evolvability (DICKINSON and SEGER 1999; WAGNER *et al.* 1999; PARTRIDGE and BARTON 2000; BROOKFIELD 2001; PAL 2001; MEIKLEJOHN and HARTL 2002; RUDEN *et al.* 2003). Any costs of an evolutionary capacitor are borne immediately, while benefits typically lie in the future. Based largely on work on mutator alleles, the general consensus among population geneticists is that it is very difficult for natural selection to favor an evolvability allele (SNIEGOWSKI *et al.* 2000).

Introducing variation via a capacitor has several advantages over introducing variation via mutation (MASEL and BERGMAN 2003). First, although both revealed variation and the revealing state of the capacitor are heritable, the entire deleterious load can typically be reversed by a single change such as the disappearance of the [*PSI*⁺] prion or the restoration of Hsp90 function. This means that if the original environment is restored, reversion to the original phenotype is straightforward. If a

¹Address for correspondence: Department of Ecology and Evolutionary Biology, University of Arizona, 1041 E. Lowell St., P.O. Box 210088, Tucson, AZ 85721. E-mail: masel@u.arizona.edu

subset of the revealed variation is adaptive, then this subset tends to acquire other mutations and to lose its dependence on the revealing mechanism. This phenomenon is known as genetic assimilation and has been well documented (WADDINGTON 1953; WADDINGTON 1956; RUTHERFORD and LINDQUIST 1998; SOLLARS *et al.* 2003; TRUE *et al.* 2004). Once genetic assimilation has occurred, the revealing mechanism can disappear, and an adaptive phenotype will have fixed in the new environment, with a very low penalty in terms of long-term deleterious load.

Second, adaptive alleles tend to become quickly separated from mutator alleles by recombination (JOHNSON 1999). In contrast, the adaptive expression of revealed alleles remains dependent on the continued action of the revelation mechanism for some time. This delay before the completion of the genetic assimilation process means that modifier alleles promoting revelation are better able to hitchhike on the basis of the adaptive alleles they reveal.

It is highly controversial, however, whether evolutionary capacitors have evolved for the purpose of promoting evolvability or whether their evolutionary properties are a mere accident, a byproduct of other functions. A previous study showed that the ability to reveal variation can increase mean fitness in an infinite population (ESHEL and MATESSI 1998), thus circumventing "neutral confinement" when genetic canalization reduces available variation to the point of evolutionary stasis (ANCEL and FONTANA 2000). This study did not, however, address whether selection for a revelation mechanism was strong enough to overcome drift in a finite population. The study therefore suggests that revelation mechanisms may be a product of natural selection, but remains inconclusive.

Another study deduced from the specific parameters associated with $[PSI^+]$ appearance that the capacity for $[PSI^+]$ formation is probably the consequence of natural selection for its evolvability properties rather than a chance event (MASEL and BERGMAN 2003). This study developed stochastic methods to study revelation in a finite population, but addressed only the specific question of the capacity for $[PSI^+]$ formation in yeast. Here I use a similar mathematical model to that used in the $[PSI^+]$ study to study the far more general question of predicting the evolution of an arbitrary revelation mechanism in a finite population.

Although the model is general to any capacitor, it is motivated by the specific biology of the $[PSI^+]$ system. We assume that cryptic genetic variation is an automatic consequence of evolution. In the context of $[PSI^+]$, cryptic genetic variation automatically builds up in untranslated DNA sequences beyond stop codons. We consider the evolutionary dynamics of revelation by examining the invasion properties of a modifier that phenotypically reveals genetic variation with a certain probability. In the context of $[PSI^+]$, variation is revealed when the

$[PSI^+]$ prion appears and translation termination is impaired as a result. A number of genetic factors influence the frequency of $[PSI^+]$ appearance, all of which are candidates for a modifier locus.

Obviously, when the population is well adapted to a constant environment, the optimal modifier never reveals variation. Environmental change events create the opportunity for the invasion of a modifier that reveals variation. The optimal probability of revelation is expected to be a function of the frequency of environmental change, the size of the population, and the extent of the adaptive advantage of the revealed variation.

OUTLINE OF MODEL

We model the environment as usually constant, punctuated by occasional events during which the revelation of variation is adaptive with fitness $1 + s$. The frequency of these events is given as θ events per generation, where θ is expected to be low. While the environment is constant, the revelation of variation is assumed to be highly deleterious. To simplify the model, we assume that individuals with revealed variation never reproduce in the absence of an environmental change event. A modifier that reveals variation with a probability of m per replication therefore has a fitness of $\sim 1 - m$ in the constant environment. This fitness is used to calculate the fixation probability and sojourn times of the modifier in the absence of environmental change events. Environmental change events are assumed to be distributed as a Poisson process. This allows the calculation of the probability that an environmental change event occurs during the sojourn time of a given modifier.

Consider a haploid population of size N that is homogeneous for an m_1 allele. A single m_2 allele appears in this population by mutation or migration. We calculate the probability that this m_2 allele successfully invades the population. In the original, constant environment, the population is well adapted. Both m_1 and m_2 alleles are occasionally at a disadvantage when variation is revealed. A Moran model is used to calculate population genetic quantities in the presence of these disadvantages. Let the probability that m_2 fixes in the absence of environmental change be p_{drift} . Let the sojourn time during which there are i m_2 alleles be τ_{1i} and the total sojourn time be τ_1 . Let the probability that environmental change occurs during the sojourn time of the m_2 allele be p_{env} and the probability conditional on m_2 eventually becoming fixed by drift be p_{env}^* .

Now consider events occurring after an environmental change event that makes the revelation of variation adaptive. At the time of environmental change, there are i m_2 alleles. If no individuals have adaptive variation revealed at this time, neutral drift takes place until an adaptive variant appears. Let x_i be the probability that no adaptive variant either already exists at the time of environmental change or appears before fixation or

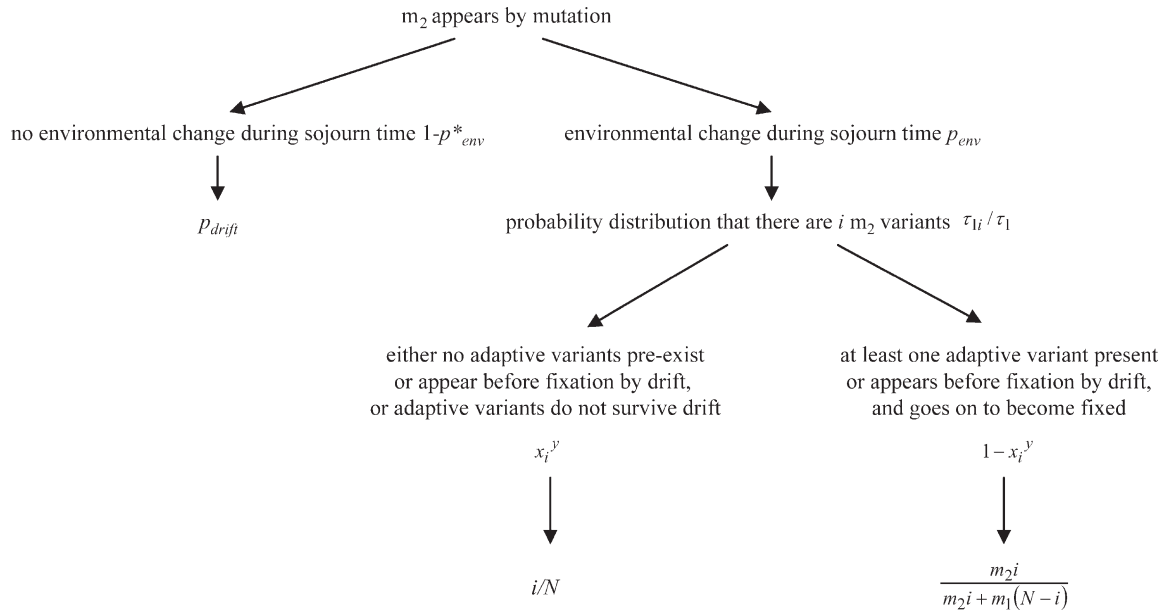


FIGURE 1.—Outline of the model used to calculate the probability that a single m_2 mutant will invade an m_1 population. The first branch considers whether environmental change occurs during the sojourn time of the m_2 allele. If not, then m_2 invasion is a matter of random drift. If environmental change does occur, then we consider the possibility that there are i m_2 individuals at the moment of environmental change for each value of $i = 1, 2, \dots, N - 1$. For each of these scenarios, we consider whether an adaptive variant is revealed by either an m_1 or an m_2 capacitor allele and survives initial stochastic processes to take over the population. If not, then which allele wins is in proportion to initial allele frequencies, and m_2 fixes with probability i/N . If revelation does lead to the fixation of an adaptation, then which allele wins is in proportion to initial rates of revelation, and m_2 fixes with probability $m_2 i / (m_2 i + m_1 (N - i))$.

extinction of m_2 by drift is complete, given an initial population of i m_2 individuals. Let y be the probability that a single adaptive variant with fitness $1 + s$ survives initial stochastic events become fixed. To calculate y , we ignore the rare event in which the capacitor switches off and the progeny of an adaptive individual loses the adaptation.

As an approximation, assume that the number of adaptive variants preexisting or appearing is distributed according to a Poisson process, and so the probability that there are no adaptive variants is $x_i = e^{-\mu_i}$, where μ_i is the mean number of adaptive variants. The probability that at least one adaptive variant will both appear and become fixed is therefore given by

$$\sum_{i=1}^{\infty} \frac{\mu_i^i}{i! e^{\mu_i}} (1 - (1 - y)^i) = 1 - e^{-\mu_i y} = 1 - x_i^y.$$

From the sketch of the possibilities shown in Figure 1, we see that the probability that m_2 invades an m_1 population is given by

$$p_{\text{invasion}}(m_1, m_2, N, \theta, s) = (1 - p_{\text{env}}^*) p_{\text{drift}} + p_{\text{env}} \times \sum_{i=1}^{N-1} \frac{\tau_i}{\tau_1} \left((1 - x_i^y) \frac{m_2 i}{m_2 i + m_1 (N - i)} + x_i^y \frac{i}{N} \right). \tag{1}$$

The terms defined above and used in Equation 1 are calculated in the APPENDIX, using the Moran model from population genetics as well as other mathematical

techniques. Due to computational constraints, an approximate solution is derived for high values of N .

If the probability that m_2 invades an m_1 population is greater than the probability with which m_1 invades an m_2 population, we conclude that m_2 is superior and will generally invade. We look for a value m_{opt} such that

$$p_{\text{invasion}}(m_{\text{opt}}, m_2, N, \theta, s) < p_{\text{invasion}}(m_2, m_{\text{opt}}, N, \theta, s)$$

for all values of m_2 . This value of m_{opt} is a stochastic extension to the concept of an evolutionary stable strategy (ESS). We then search through possible values of m to calculate the optimal rate of revealing variation m_{opt} as a function of the parameters N , θ , and s .

RESULTS

In Figure 2A we show the optimal rate of revelation m_{opt} as a function of the three parameters of the model, N , θ , and s , using the approximate solution. For large $N > 1/\theta$, we get $m_{\text{opt}} \approx 1/N$. For small N below some cutoff value, m_{opt} declines rapidly toward zero. In between these two extremes, m_{opt} reaches a plateau value of $\sim \theta$.

The only effect of the parameter s is on the minimum cutoff value of N . Very small values of $s < 1/N$ indicate such weak selection as to be barely appreciable. For substantive values of $s \gg 1/N$, the precise value of s does not affect the results.

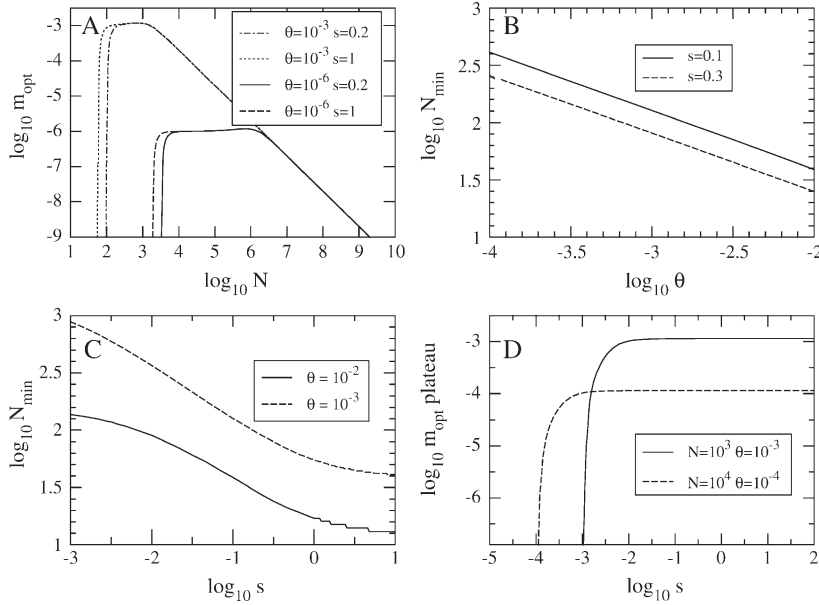


FIGURE 2.—The optimal rate of revelation m_{opt} . (A) Above a certain minimum population size, m_{opt} is approximately given by either the environmental change rate θ or $1/N$, whichever is smaller. (B and C) The minimum population size N_{min} for selection for capacitance to overcome random drift as a function of s and θ . (D) The value of m_{opt} with $N = 1/\theta$.

In Figure 2, B and C, we look at the minimum value of N more carefully as a function of θ and s , showing the smallest value of N for which $m_{\text{opt}} > 0$. We can see from the gradient in Figure 2B that N_{min} increases with $\theta^{-1/2}$. N_{min} also increases in a more complex way with decreasing s , as shown in Figure 2C.

In Figure 2D we examine the value of m_{opt} around the plateau by taking $N = 1/\theta$. The plateau value of m_{opt} is approximately equal to θ as long as $s > 1/N$ and decreases dramatically for smaller values of s such that positive selection is slight or negligible.

In summary, for a population size N above a minimum cutoff, m_{opt} is approximately given by either $1/N$ or θ , whichever is smaller. Unicellular eukaryotes such as yeast typically have an effective population size on the order of 10^7 – 10^8 (LYNCH and CONERY 2003). Since it seems plausible that environmental change events are more common than once every 10^7 generations, we predict that m_{opt} should be given by $\sim 1/N$. The prion [PSI⁺] acts as a revelation mechanism of the type modeled in this article, and [PSI⁺] appears in the yeast *Saccharomyces cerevisiae* at a rate of $\sim 10^{-5}$ – 10^{-7} per replication (LUND and COX 1981). This rate of appearance is in approximate agreement with the model's prediction, allowing for a wide margin of error for both the estimation of the effective population size and the estimation of the [PSI⁺] appearance rate.

Looking more generally at other organisms, we find that the evolution of capacitance is possible even for populations with small population sizes (e.g., 10^3) given a fairly moderate minimum frequency of environmental change (e.g., 10^{-4}), as shown in Figure 2, B and C.

In Figure 2 we explored the value of the optimal rate of revelation m_{opt} . In Figure 3, we examine the degree to which the optimal rate is favored over alternatives. This is shown by the frequency with which an allele invades and is invaded by the ESS. When new individuals

with different values of m appear by mutation, then the frequency of allele substitution is related to the quantity Np_{invasion} . In Figure 3, A and B, we show the kinetics of invasion of a plateau value of m_{opt} when $N = 1/\theta$. The difference between the rates of invasion and counterinvasion gives an indication of the effective selective advantage of an m_{opt} allele. It can be seen that these kinetics are broadly similar for different values of N and s , respectively. High rates of revelation will readily be invaded by m_{opt} . Low rates of revelation are under comparatively mild selection, becoming stronger with large population sizes and larger potential adaptations.

In Figure 3C we show the invasion kinetics for other values of $N \neq 1/\theta$. The dotted lines show a smaller value of N , still on the plateau but closer to the cutoff minimum N . The dashed lines show a value of $N > 1/\theta$, for which m_{opt} is lower. Large population sizes lead to a greater selective difference for low values of m .

Note that around m_{opt} both the invasion rate and the counterinvasion rate are faster than that for a neutral allele. This is because selection is working at two different timescales. A lower value of m invades often during the periods between environmental change events. A higher value of m regularly invades around the time of environmental change. The total flux of alleles through the system is therefore higher than that for a static environment.

In Figure 4 we focus on the effect of population size on the rates of invasion and counterinvasion. In Figure 4, A and B, we look at the rates of invasion and counterinvasion by which m_{opt} invades $m_{\text{opt}}/100$ (Figure 4A) and $100m_{\text{opt}}$ (Figure 4B). We see that large $N \gg \theta$ leads to a greater flux of alleles with both invasion and counterinvasion occurring with probability $> 1/N$. Curves are close to superimposable for a constant value of $N\theta > 1$ for different values of θ .

In Figure 4, C and D, we look at the ratio of the rates

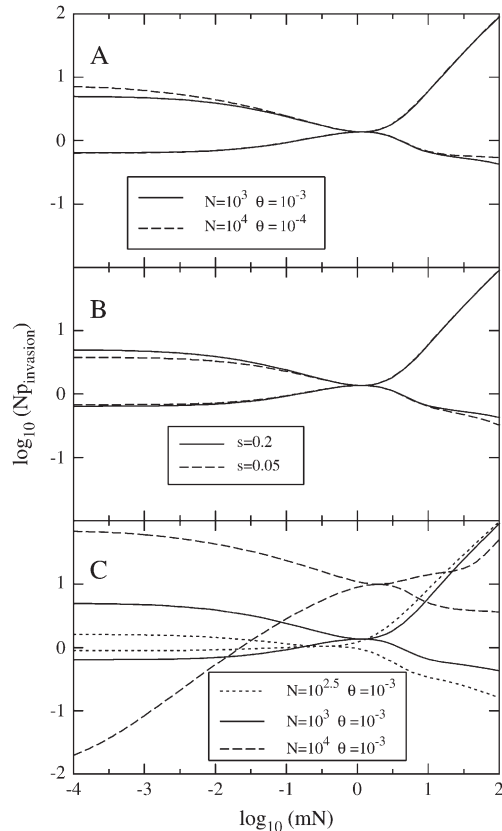


FIGURE 3.—The extent to which the optimal rate of revelation is favored over alternatives. Both the rate at which m_{opt} invades and the rate at which it is invaded are shown. Where not specified, $s = 0.2$, $N = 10^3$, and $\theta = 10^{-3}$. (A and B) Around $N = 1/\theta$, results are relatively insensitive to parameter values. Selection against revelation more frequent than the optimum is stronger than selection against revelation less frequent than the optimum. (C) Other values of $N \neq 1/\theta$. Selection against infrequent revelation is much stronger for large N .

of invasion and counterinvasion. We see in Figure 4C that the advantage of m_{opt} over smaller values of m increases with population size before reaching a plateau at some value of $N > 1/\theta$. This plateau ratio of invasion to counterinvasion is given approximately by the ratio of m values. In Figure 4D we see that the advantage of m_{opt} over larger values of m is greatest for $N \approx 1/\theta$. It falls to a plateau for $N\theta$ greater than the ratio of m values.

In Figure 4E we examine the extent of the turnover rate of similar alleles by looking at the rate that an allele of $0.99m_{\text{opt}}$ can invade m_{opt} . We see that flux is neutral when $N < 1/\theta$, and then for larger values of N flux is approximately proportional to $1/(N\theta)$.

In Figures 5 and 6 we test the accuracy of the approximate solution. This is computationally possible only for small N . Figure 5 is the equivalent of Figure 2A, and Figure 6 roughly corresponds to Figure 3C. It can be seen that the approximation does not have a large effect on the results. In Figure 5 we see that there is no true plateau value of m_{opt} and that the approximation overestimates m_{opt} in the range $N_{\text{min}} < N < 1/\theta$. In Figure 6

we see that the approximation overestimates selection against rates of revelation lower than m_{opt} . Generally, however, the difference between the exact and approximate solutions is not large, and overall trends are maintained.

In summary, evolution toward m_{opt} occurs most rapidly for large $N \gg 1/\theta$. The overall advantage of optimal revelation over less frequent revelation is also greatest for $N \gg 1/\theta$. The advantage over more frequent capacitance, in contrast, is weaker at large N and most efficient for $N \approx 1/\theta$. For small values of $N < 1/\theta$, the evolution of capacitance may in principle be favored, but the effect may be too marginal so that, combined with the less frequent appearance of capacitor alleles to begin with, the evolution of capacitance is far from certain. In contrast, the evolution of capacitance in large populations with $N > 1/\theta$ seems highly likely.

DISCUSSION

Interpretation of the parameter θ : Here we have assumed that a reservoir of latent variation always exists. In other words, cryptic genetic variation is an automatic consequence of evolution, extrinsic to the evolvability model of revelation presented here.

Nevertheless, the amount of potentially useful phenotypic variation may vary substantially according to the population genetic and demographic history of the population. All these effects are collapsed into the single parameter θ , which represents the probability that a revelation event is adaptive and that this adaptation cannot be achieved by tapping standing variation in the population. The model could be extended to cover factors such as the level of preadaptation (ESHEL and MATESSI 1998) by studying their effect on the parameter θ . As another example, the time between successive environmental change events may be important in determining the amount of latent variation that has had time to build up before being purged again. This effect, too, could be modeled by allowing θ to be a function of the time since the last selective sweep.

Note that the molecular nature of the revealed variation may vary according to which capacitor mechanism is involved. For example, $[PSI^+]$ reveals genetic variation beyond stop codons while Hsp90 reveals underlying weaknesses in complex networks. Models seeking to break θ into components may need to focus on a particular capacitor and the specific nature of the variation it reveals. The model presented here can be seen as an overall framework that can be applied to any capacitor system.

Stress response: We have assumed that the rate m of revelation is constant. In practice, for a wide range of mechanisms, stress increases the rate at which variation is revealed (BADYAEV 2005). This means that m is higher at a time of environmental change, increasing positive selection for capacitance, while m is lower in a static environment, decreasing negative selection against ca-

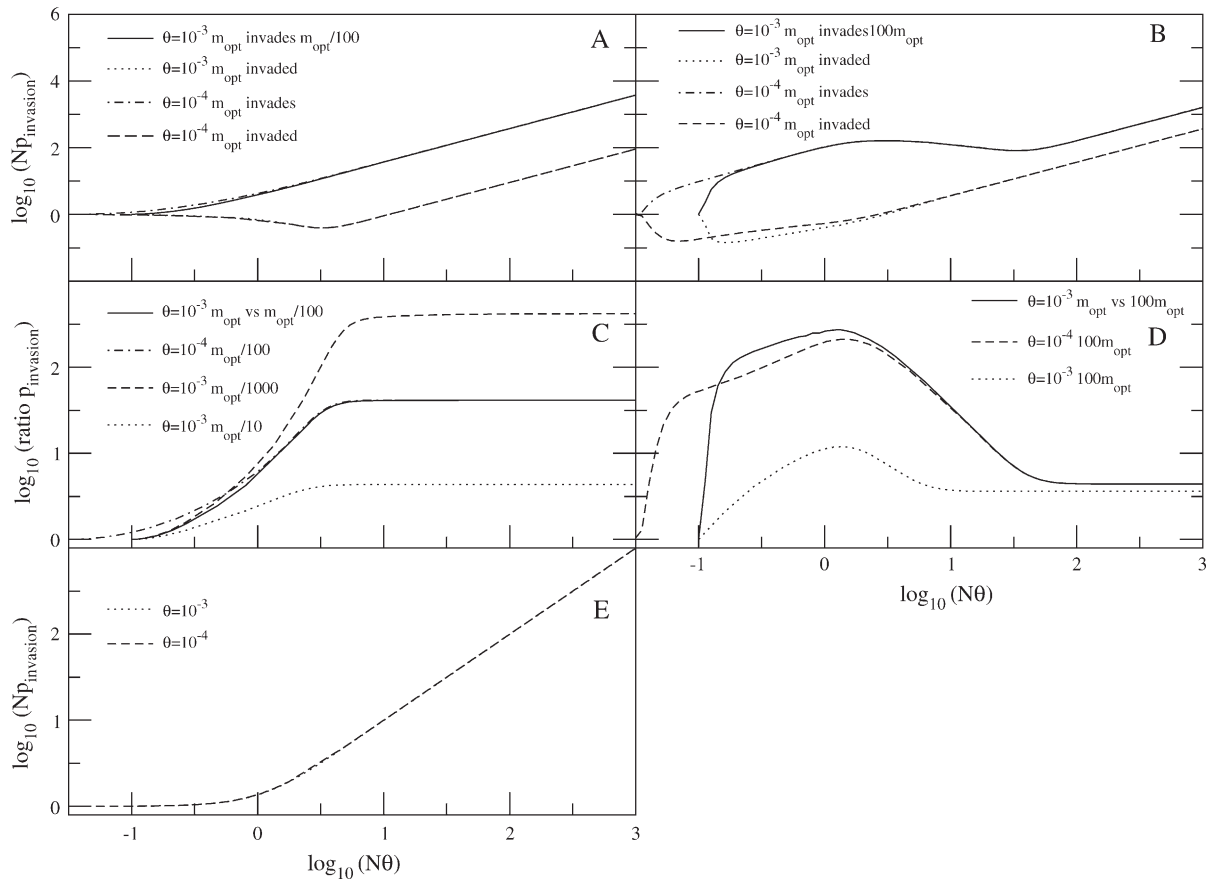


FIGURE 4.—The effect of population size on the extent to which the optimal rate of revelation is favored over alternatives. Population size varies over the x -axis. $s = 0.2$ throughout. (A) Rate of invasion and counterinvasion of m_{opt} vs. $m_{\text{opt}}/100$. Both rates increase with large population sizes, leading to a high flux of alleles through the system. (B) Rate of invasion and counterinvasion of m_{opt} vs. $100m_{\text{opt}}$. (C and D) The ratio of invasion and counterinvasion rates gives a single measure of the strength of selection. (E) The flux of alleles, as measured by the rate with which a $0.99m_{\text{opt}}$ allele invades m_{opt} , depends on the product $N\theta$.

capacitance. Clearly, the assumption of constant m is conservative in estimating the extent to which natural selection favors capacitance.

The evolution of capacitance can be seen as a two-stage process. In the first stage, an evolutionary capacitor mechanism arises with a constant rate of revelation m . This is the stage described by the present model. In the second stage, the capacitor acquires regulatory features that allow it to specifically respond to stress. A single parameter m can be seen as a special case of a more general model in which two revelation rates represent the presence and absence of stress. The extension of the model presented here to incorporate this additional parameter is straightforward.

This two-stage scenario is clearly relevant to the biology of $[PSI^+]$, where prion appearance may have been random at first, until modifiers evolved to regulate it. In contrast, revelation via Hsp90 depletion may be a direct consequence of stress, and hence a stress response may have been an intrinsic part of the capacitor system from the beginning.

Plausibility of selection on $[PSI^+]$ and Hsp90-mediated revelation: Many factors are known to influence $[PSI^+]$ appearance (CHERNOFF 2001), and there have been ample evolutionary time (CHERNOFF *et al.* 2000; KUSHNIROV *et al.* 2000; SANTOSO *et al.* 2000; NAKAYASHIKI *et al.* 2001) and sufficient population size for selection on these potential modifiers to have occurred. In addition, it is possible to measure the frequency of $[PSI^+]$ appearance under controlled conditions (LUND and COX 1981). As we discover more about the modifiers of $[PSI^+]$ appearance, it should become possible to search for evidence of the action of selection on these modifiers.

Hsp90-mediated revelation has been observed in both *Drosophila* (RUTHERFORD and LINDQUIST 1998) and *Arabidopsis* (QUEITSCH *et al.* 2002), so again there have also been ample evolutionary time and sufficient population size ($\sim 10^6$) for selection on revelation modifiers to have occurred. The mechanism of Hsp90-mediated revelation and the nature of possible modifiers are not yet well understood, however (SOLLARS *et al.* 2003). Also,

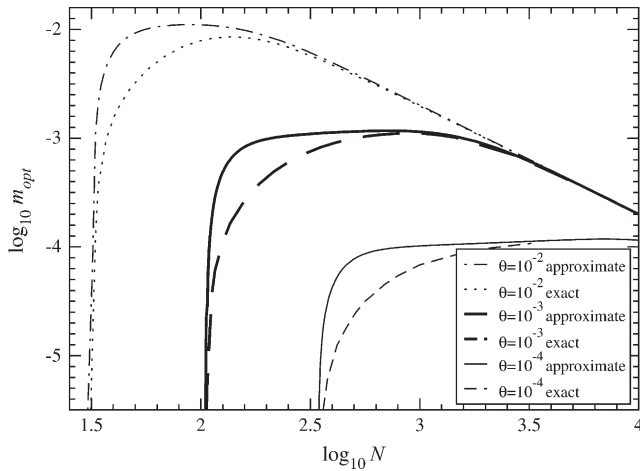


FIGURE 5.—The optimal rate of revelation m_{opt} given by the exact solution *vs.* the approximate solution. For large N , the two converge. N_{min} is the same for both solutions. The greatest difference is that the exact solution gives smaller values of m_{opt} in the range $N_{min} < N < 1/\theta$. $s = 0.2$ throughout.

Hsp90-mediated revelation has not been seen in a natural setting, but has instead been hypothesized on the basis of laboratory results on the partial loss of function of Hsp90 due to either mutation or geldanamycin treatment. The same cryptic genetic variation can sometimes be revealed by increased temperature alone (QUEITSCH *et al.* 2002), but the involvement of Hsp90 in this revelation has not been proven. Until the molecular details of Hsp90-mediated revelation are better understood, it will be hard to assess the action of selection on them.

Comparison with mutator models: It is instructive to compare the capacitor model developed here to mutator models. In the capacitor model, beneficial variants appear on a stochastic basis. Since most mutator models treat beneficial mutations deterministically, I focus the comparison on a recent mutator model that also included stochasticity in the appearance of beneficial mutations (JOHNSON 1999).

The capacitor model differs from this mutator model in several ways. First, the capacitor model also treats deleterious revelation events according to a stochastic rather than a deterministic process. The stochasticity is captured in the sojourn times derived from a Markov chain model and was made possible by the simplifying assumption that deleterious variants never reproduce. Classical mutations are typically mildly deleterious and handicap the organism in the form of an accumulated deleterious mutation load at multiple sites (JOHNSON 1999). In contrast, capacitors reveal multiple and therefore significant deleterious variation at a single stroke. Although the capacitor model could formally be seen as a novel model of a mutator allele, the assumption of rare and strong deleterious effects is more suited to the biology of a capacitor.

Second, the mutator model assumes that beneficial

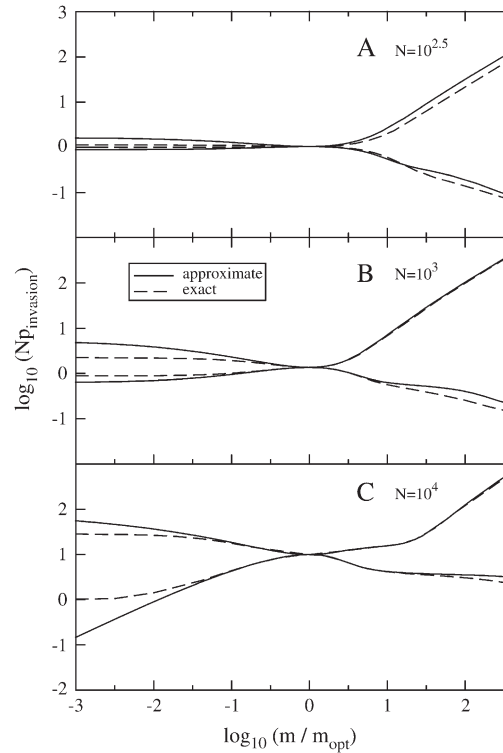


FIGURE 6.—The extent to which the optimal rate of revelation is favored over alternatives given by the exact solution *vs.* the approximate solution. The approximation slightly exaggerates the selective advantage of m_{opt} over less frequent revelation. $s = 0.2$ and $\theta = 10^{-3}$ throughout.

mutations that are destined to become fixed appear at some rate K that is independent of the population size N and the mutation rate m (JOHNSON 1999). In contrast, I assume that *opportunities* for beneficial mutations arise at a constant rate. If a beneficial variant takes longer to appear than the sojourn time of a capacitor allele, then the opportunity for a beneficial adaptation does not affect the evolution of capacitance. The rate at which the opportunity for beneficial mutations is exploited depends on both N and m .

Third, mutator models conclude that for an asexual population, the equilibrium genomic mutation rate \hat{U} should be approximately equal to the rate of beneficial substitutions K , when the metabolic cost of suppressing mutators is neglected (KIMURA 1967; JOHNSON 1999). The deleterious genomic mutation rate is estimated as $U \approx 0.1-1$ for higher eukaryotes (DRAKE *et al.* 1998) and $K \approx 0.002-0.03$ (JOHNSON 1999). Comparing these two numbers, the combined effects of selection on deleterious and beneficial mutations are likely to favor a reduction in the mutation rate toward K (JOHNSON 1999).

In the context of a capacitor, we obtain a seemingly similar result: the optimal rate of revelation m_{opt} is predicted to be no larger than the rate of environmental change θ . For the yeast prion [PSI^+], the observed rate

of revelation is of the order $m = 10^{-5}$ – 10^{-7} (LUND and COX 1981), far lower than U . This is at least partly offset by the fact that θ should be lower than K , since only some proportion of adaptive substitutions will be a consequence of the action of capacitors. For capacitance to work, a revealed variant must be so advantageous as to compensate for all other variation revealed at the same time. Nevertheless, it is easily possible that $\theta \geq m$, leaving room for natural selection to increase the rate of revelation, in stark contrast to the interpretation of a mutator model.

The capacitor model described here is more closely related to a two-locus model of selection on the mutation rate at a single, specific locus. Early work suggested that selection in this case is too weak to overcome drift and selection at linked loci. In the case of a capacitor, the “locus” is very wide ranging in its effects. We have calculated the relevant invasion and counterinvasion rates and shown that selection can be strong enough to overcome drift.

Fourth, an additional difference between a capacitor and a mutator exists in sexual populations. Mutator alleles quickly become separated by recombination from adaptive alleles that they generate (JOHNSON 1999). In contrast, a novel adaptation remains dependent on the revealing capacitor and remains dependent until genetic assimilation is complete. This makes the evolution of capacitance in a sexual system not so dissimilar to the asexual case, in stark contrast to the situation for mutator evolution. The model here is formally a model of a haploid asexual population. I argue that although some quantitative difference may be expected in a sexual population, the difference will be much smaller than that for mutators.

Note that the model presented here considers only the case when the demand for adaptation is met by a capacitor, rather than by standing variation or *de novo* mutation. Related previous work also considered alternative means of adaptation via a point mutation occurring at rate m_{point} and found that the inclusion of mutation in the model affected only the results for $m \leq m_{\text{point}}$ (MASEL and BERGMAN 2003). This means that when the model predicts a very small value for m_{opt} , adaptation may occur through mutation rather than through capacitance. In this case it is best to interpret m_{opt} as including all means of generating variance, including mutation. Capacitance evolves to increase the value of m when mutation is lower than m_{opt} . When m_{opt} gets extremely low and therefore consists primarily of mutation, then capacitance will not evolve. This becomes an issue primarily for extremely large populations.

Comparison with modeling time to adaptation: A common approach to the evolution of evolvability is to test the speed with which different genetic systems find a new optimum on a fitness landscape (HINTON and NOWLAN 1987; ANCEL 2000; HANSEN *et al.* 2000; BERGMAN

and SIEGAL 2003; BEHERA and NANJUNDIAH 2004). This approach has been used to show the advantages of an evolutionary capacitor (HANSEN *et al.* 2000; BERGMAN and SIEGAL 2003), but is unable to compare them to the costs. At best it can show that there might, under some circumstances, be an advantage to an evolutionary capacitor. It is unable to define what those circumstances are.

The approach taken here grafts two models together. One model describes a generally constant environment and calculates the costs of capacitance. The other model describes events surrounding a rare environmental change event and calculates the benefits of capacitance. The two models are combined by using a summation over all possible states of the system at the time of an environmental change event. This combined model is able to predict under what circumstances and with what properties evolutionary capacitance will evolve.

A second problem with modeling the time to adaptation is that it equates a rapid approximate adaptation followed by a slow approach to the optimum with a path of equal total length consisting of a long pause followed by a jump and rapid approach to the optimum (ANCEL 2000). When the two systems exist within a single population, the subpopulation with the rapid approximate adaptation is likely to take over before the other subpopulation has the opportunity to jump straight to the optimum. This represents the evolutionary advantage of the stopgap nature of a capacitor. In contrast to modeling the total time to adaptation, the model presented here explicitly captures the dynamics of a stopgap mechanism. In this context, revelation represents a rapid approximate adaptation, and the final optimum corresponds to the state after genetic assimilation and the reversal of the capacitor mechanism.

Finally, models of the time to adaptation are often based on numerical simulations, which restrict the range of population size that can be considered. Although the summations performed here are computationally intensive, the system is still analytical, and so realistic population sizes for organisms such as yeast can be studied.

Comparison with earlier capacitor models: Three main approaches to modeling capacitors have been taken. The first is described in the section above (BERGMAN and SIEGAL 2003). The second is based on an infinite population and the deterministic dynamics relating to mean fitness calculations (ESHEL and MATESSI 1998). The stochastic approach taken here has the advantage of predicting the evolution of capacitance in a finite population.

The third previous approach to modeling capacitors (MASEL and BERGMAN 2003) is closely related to this work, but the model presented here incorporates two significant improvements. First, the criteria for an environmental change event have been broadened. In the earlier work, an environmental change event was as-

sumed to be severe enough to lead to population extinction if variation was not revealed. In this work, any point in time in which it is advantageous to reveal variation is included as a possible event, without invoking the stronger assumption of extinction. Second, the earlier work neglected stochastic events following an environmental change event and assumed that if an adaptive variation were revealed, it would automatically take over the population. This work explicitly models stochastic events following environmental change.

We have not explicitly considered the possibility that variation may be present without resorting to a capacitor. This possibility was included in a related mathematical model (MASEL and BERGMAN 2003). If adaptive variation is available in a conventional form without resorting to revealing large quantities of presumably deleterious variation in other traits, then this conventional variation will become fixed. The environmental change event specified by the rate parameter θ should be interpreted to mean those environmental change events for which conventional variation is not sufficient for adaptation. It is possible, therefore, that θ decreases for large N .

Stochastic extension of ESS: The traditional concept of an ESS is based on the inability to be invaded by any other strategy. It is normally defined using a mean fitness approach within the context of an infinite population with deterministic dynamics (MAYNARD SMITH 1982).

In a finite population of interacting agents, the ESS may be different. A single type A individual in a population of $N - 1$ type B individuals creates an asymmetry, since the A individual encounters only B individuals, while B individuals encounter an A individual with probability $1/(N - 1)$. The fact that an individual in a finite population cannot interact with itself changes the location of the ESS and has resulted in new definitions and criteria for the ESS (RILEY 1979; MAYNARD SMITH 1988; SCHAFFER 1988; NEILL 2004).

These new definitions continue to be based on mean fitness calculations and deterministic dynamics. Stochastic fluctuations due to random drift can be incorporated into a model when mutations are rare relative to the timescale at which fixation occurs (PROULX and DAY 2001). This is achieved by using the probability of fixation instead of using mean fitness (PROULX and DAY 2001). For example, in one recent work, an ESS was defined as one that cannot be invaded by any other strategy with probability $>1/N$ (NOWAK *et al.* 2004).

I take this definition further. Around the ESS, invasion of rare revelation is favored in the constant environment and happens with a frequency $>1/N$. Invasion of more frequent revelation is favored by environmental change events and also happens with a frequency $>1/N$. I define an ESS as a strategy for which any competing strategy is counterinvaded with a greater probability than it invades.

Conclusions: The evolutionary costs and benefits of

a capacitor take place on different timescales. The benefits are felt rarely and strongly, while the costs are felt frequently and mildly. Selective pressures take the form of second-order or indirect selection. This means that although an individual may have wild-type fitness, its offspring might not. This behavior is not easily captured by a traditional selection coefficient, making it difficult to combine the selective advantages and disadvantages of the system into a single model.

The capacitor model described here is able to calculate the impact of extreme but rare events on the nature of evolution. These features are by nature almost impossible to capture experimentally. Here I have shown that taking into account both costs and benefits, evolutionary capacitance may be favored by natural selection. This was done using a fully stochastic, analytical model of a finite population. For populations larger than a minimum cutoff size the optimal revelation rate is approximately given by either the inverse of the population size or the frequency with which revelation is adaptive, whichever is smaller. The degree to which the optimum is favored increases with the population size. Capacitance is expected to evolve under a range of biologically realistic conditions, particularly in large populations.

I thank Aviv Bergman, Marc Feldman, Lilach Hadany, Heather Maughan, Mark Siegal, and Emile Zuckerkandl for helpful discussions and careful reading of the article. I acknowledge National Institutes of Health grant GM28016 to M. W. Feldman and the BIO5 Institute for financial support.

LITERATURE CITED

- ANCEL, L. W., 2000 Undermining the Baldwin expediting effect: Does phenotypic plasticity accelerate evolution? *Theor. Popul. Biol.* **58**: 307–319.
- ANCEL, L. W., and W. FONTANA, 2000 Plasticity, evolvability, and modularity in RNA. *J. Exp. Zool.* **288**: 242–283.
- BADYAEV, A. V., 2005 Stress-induced variation in evolution: from behavioral plasticity to genetic assimilation. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **272**: 877–886.
- BEHERA, N., and V. NANJUNDIAH, 2004 Phenotypic plasticity can potentiate rapid evolutionary change. *J. Theor. Biol.* **226**: 177–184.
- BERGMAN, A., and M. L. SIEGAL, 2003 Evolutionary capacitance as a general feature of complex gene networks. *Nature* **424**: 549–552.
- BROOKFIELD, J. F. Y., 2001 Evolution: the evolvability enigma. *Curr. Biol.* **11**: R106–R108.
- CHERNOFF, Y. O., 2001 Mutation processes at the protein level: Is Lamarck back? *Mutat. Res.* **488**: 39–64.
- CHERNOFF, Y. O., A. P. GALKIN, E. LEWITIN, T. A. CHERNOVA, G. P. NEWNAM *et al.*, 2000 Evolutionary conservation of prion-forming abilities of the yeast Sup35 protein. *Mol. Microbiol.* **35**: 865–876.
- DE VISSER, J., J. HERMISSON, G. P. WAGNER, L. A. MEYERS, H. BAGHERI-CHAICHIAN *et al.*, 2003 Perspective: evolution and detection of genetic robustness. *Evolution* **57**: 1959–1972.
- DICKINSON, W. J., and J. SEGER, 1999 Cause and effect in evolution. *Nature* **399**: 30.
- DRAKE, J. W., B. CHARLESWORTH, D. CHARLESWORTH and J. F. CROW, 1998 Rates of spontaneous mutation. *Genetics* **148**: 1667–1686.
- ESHEL, I., and C. MATESSI, 1998 Canalization, genetic assimilation and preadaptation: a quantitative genetic model. *Genetics* **149**: 2119–2133.
- EWENS, W. J., 1979 *Mathematical Population Genetics*. Springer, Berlin/Heidelberg, Germany.

- HANSEN, T. F., A. J. R. CARTER and C. H. CHIU, 2000 Gene conversion may aid adaptive peak shifts. *J. Theor. Biol.* **207**: 495–511.
- HINTON, G. E., and S. J. NOWLAN, 1987 How learning can guide evolution. *Complex Syst.* **1**: 495–502.
- IWASA, Y., F. MICHOR and M. A. NOWAK, 2004 Stochastic tunnels in evolutionary dynamics. *Genetics* **166**: 1571–1579.
- JOHNSON, T., 1999 Beneficial mutations, hitchhiking and the evolution of mutation rates in sexual populations. *Genetics* **151**: 1621–1631.
- KIMURA, M., 1967 On the evolutionary adjustment of spontaneous mutation rates. *Genet. Res.* **9**: 23–34.
- KRAKAUER, D. C., and J. B. PLOTKIN, 2002 Redundancy, antiredundancy, and the robustness of genomes. *Proc. Natl. Acad. Sci. USA* **99**: 1405–1409.
- KUSHNIROV, V. V., N. KOCHNEVA-PERVUKHOVA, M. B. CHECHENOVA, N. S. FROLOVA and M. D. TER-AVANESYAN, 2000 Prion properties of the Sup35 protein of yeast *Pichia methanolica*. *EMBO J.* **19**: 324–331.
- LUND, P. M., and B. S. COX, 1981 Reversion analysis of [*psi*⁻] mutations in *Saccharomyces cerevisiae*. *Genet. Res.* **37**: 173–182.
- LYNCH, M., and J. S. CONERY, 2003 The origins of genome complexity. *Science* **302**: 1401–1404.
- MASEL, J., and A. BERGMAN, 2003 The evolution of the evolvability properties of the yeast prion [PSI⁺]. *Evolution* **57**: 1498–1512.
- MAYNARD SMITH, J., 1982 *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK.
- MAYNARD SMITH, J., 1988 Can a mixed strategy be stable in a finite population? *J. Theor. Biol.* **130**: 247–251.
- MEIKLEJOHN, C. D., and D. L. HARTL, 2002 A single mode of canalization. *Trends Ecol. Evol.* **17**: 468–473.
- NAKAYASHIKI, T., K. EBIHARA, H. BANNAI and Y. NAKAMURA, 2001 Yeast [PSI⁺] “prions” that are cross-transmissible and susceptible beyond a species barrier through a quasi-prion state. *Mol. Cell* **7**: 1121–1130.
- NEILL, D. B., 2004 Evolutionary stability for large populations. *J. Theor. Biol.* **227**: 397–401.
- NOWAK, M. A., M. C. BOERLIJST, J. COOKE and J. M. SMITH, 1997 Evolution of genetic redundancy. *Nature* **388**: 167–171.
- NOWAK, M. A., A. SASAKI, C. TAYLOR and D. FUDENBERG, 2004 Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**: 646–650.
- PAL, C., 2001 Yeast prions and evolvability. *Trends Genet.* **17**: 167–169.
- PARTRIDGE, L., and N. H. BARTON, 2000 Evolving evolvability. *Nature* **407**: 457–458.
- PROULX, S. R., and T. DAY, 2001 What can invasion analyses tell us about evolution under stochasticity in finite populations? *Selection* **1–2**: 1–15.
- QUEITSCH, C., T. A. SANGSTER and S. LINDQUIST, 2002 Hsp90 as a capacitor of phenotypic variation. *Nature* **417**: 618–624.
- RICE, S. H., 1998 The evolution of canalization and the breaking of von Baer’s laws: modeling the evolution of development with epistasis. *Evolution* **52**: 647–656.
- RICE, S. H., 2002 A general population genetic theory for the evolution of developmental interactions. *Proc. Natl. Acad. Sci. USA* **99**: 15518–15523.
- RILEY, J. G., 1979 Evolutionary equilibrium strategies. *J. Theor. Biol.* **76**: 109–123.
- RUDEN, D. M., M. D. GARFINKEL, V. E. SOLLARS and X. Y. LU, 2003 Waddington’s widge: Hsp90 and the inheritance of acquired characters. *Semin. Cell Dev. Biol.* **14**: 301–310.
- RUTHERFORD, S. L., and S. LINDQUIST, 1998 Hsp90 as a capacitor for morphological evolution. *Nature* **396**: 336–342.
- SANTOSO, A., P. CHIEN, L. Z. OSHEROVICH and J. S. WEISSMAN, 2000 Molecular basis of a yeast prion species barrier. *Cell* **100**: 277–288.
- SCHAFFER, M. E., 1988 Evolutionarily stable strategies for a finite population and a variable contest size. *J. Theor. Biol.* **132**: 469–478.
- SNIEGOWSKI, P. D., P. J. GERRISH, T. JOHNSON and A. SHAVER, 2000 The evolution of mutation rates: separating causes from consequences. *BioEssays* **22**: 1057–1066.
- SOLLARS, V., X. Y. LU, L. XIAO, X. Y. WANG, M. D. GARFINKEL *et al.*, 2003 Evidence for an epigenetic mechanism by which Hsp90 acts as a capacitor for morphological evolution. *Nat. Genet.* **33**: 70–74.
- TRUE, H. L., and S. L. LINDQUIST, 2000 A yeast prion provides a mechanism for genetic variation and phenotypic diversity. *Nature* **407**: 477–483.
- TRUE, H. L., I. BERLIN and S. L. LINDQUIST, 2004 Epigenetic regulation of translation reveals hidden genetic variation to produce complex traits. *Nature* **431**: 184–187.
- VAN NIMWEGEN, E., J. P. CRUTCHFIELD and M. HUYNEN, 1999 Neutral evolution of mutational robustness. *Proc. Natl. Acad. Sci. USA* **96**: 9716–9720.
- WADDINGTON, C. H., 1953 Genetic assimilation of an acquired character. *Evolution* **7**: 118–126.
- WADDINGTON, C. H., 1956 Genetic assimilation of the bithorax phenotype. *Evolution* **10**: 1–13.
- WAGNER, A., 1996 Does evolutionary plasticity evolve? *Evolution* **50**: 1008–1023.
- WAGNER, G. P., C. H. CHIU and T. F. HANSEN, 1999 Is Hsp90 a regulator of evolvability? *J. Exp. Zool.* **285**: 116–118.
- WILKE, C. O., J. L. WANG, C. OFRIA, R. E. LENSKI and C. ADAMI, 2001 Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* **412**: 331–333.

Communicating editor: G. GIBSON

APPENDIX

Before environmental change: Consider a point in time before an environmental change event, when there are i m_2 alleles and $N - i$ m_1 alleles. At each time step, one individual is chosen at random to reproduce and another to die. These individuals may be the same. It is assumed that individuals express variation with probability m_1 or m_2 depending on their genotype and that those individuals expressing variation do not reproduce. The probability that the next individual chosen to reproduce is m_2 is therefore given by

$$\frac{(1 - m_2) i}{(1 - m_1)(N - i) + (1 - m_2) i}.$$

The probability that the next individual chosen to die is m_1 is given by $(N - i)/N$. The probability that the number of m_2 variants increases from i by one is then given by the probability that an m_2 individual is chosen to reproduce while an m_1 individual is chosen to die:

$$\lambda_i = \frac{(1 - m_2) i (N - i)}{((1 - m_1)(N - i) + (1 - m_2) i) N}.$$

The probability that the number of m_2 variants decreases from i by one is given by the probability that an m_1 individual is chosen to reproduce while an m_2 individual is chosen to die:

$$\mu_i = \frac{(1 - m_1)(N - i) i}{((1 - m_1)(N - i) + (1 - m_2) i) N}.$$

Following (EWENS 1979), define

$$\rho_0 = 1, \rho_i = \frac{\prod_{j=1}^i \mu_j}{\prod_{j=1}^i \lambda_j} = \left(\frac{1 - m_1}{1 - m_2} \right)^i.$$

The probability of fixation by drift starting from i individuals is then

$$\pi_i = \frac{\rho_i - 1}{\rho_N - 1}$$

(EWENS 1979), and the probability of fixation by drift starting from a single variant individual is

$$p_{\text{drift}} = \frac{\rho_1 - 1}{\rho_N - 1}. \tag{A1}$$

Then the sojourn time τ_{1i} during which there are i descendents of a single original variant is given by

$$\tau_{1i} = \begin{cases} \frac{(1 + (m_2 - m_1)/(\rho_N - 1)(1 - m_2))(N - m_2 N + m_2 - m_1)}{(1 - m_2)(N - 1)}, & i = 1 \\ \frac{(\rho_{N-i} - 1)(N - m_2 N + m_2 i - m_1 i)}{(1 - m_1) i (N - i) (\rho_N - 1)}, & i = 2, \dots, N - 1 \end{cases} \tag{A2}$$

(EWENS 1979), where the unit of time is one generation or N rounds in the Moran model. The total sojourn time before extinction or fixation is

$$\tau_1 = \sum_{i=1}^{N-1} \tau_{1i}, \tag{A3}$$

and the probability that there will be i descendents at a particular time, given that neither extinction nor fixation has yet occurred, is given by τ_{1i}/τ_1 . The sojourn times conditional on fixation rather than elimination by drift ultimately occurring are given by

$$\tau_{1i}^* = \begin{cases} \tau_{1i} \frac{\pi_i}{\pi_1} = \tau_{1i}, & i = 1 \\ \frac{(\rho_i - 1)(\rho_{N-i} - 1)(N - m_2 N + m_2 i - m_1 i)}{(\rho_1 - 1)(1 - m_1) i (N - i) (\rho_N - 1)}, & i = 2, \dots, N - 1 \end{cases}$$

$$\tau_1^* = \sum_{i=1}^{N-1} \tau_{1i}^*$$

(EWENS 1979).

Consider a newly appeared m_2 mutant. Let the mean frequency of a certain class of environmental change be θ events per generation, and assume that these events occur according to a Poisson process. The probability that environmental change happens at least once during the sojourn time of the m_2 lineage is then

$$p_{\text{env}} = 1 - e^{-\theta \tau_1}. \tag{A4}$$

The probability of environmental change occurring during the sojourn time, conditional on fixation by drift ultimately occurring, is given by

$$p_{\text{env}}^* = 1 - e^{-\theta \tau_1^*}. \tag{A5}$$

After environmental change: Consider the situation after environmental change in which there are i adapted individuals. The probability that an adaptive individual is chosen to reproduce is

$$\frac{i(1 + s)}{i(1 + s) + N - i}$$

and the probability that an adaptive individual is chosen to die is i/N . Then we derive the probability that a single adaptive variant becomes fixed as

$$y = \frac{1 - \rho_1}{1 - \rho_N}, \quad \text{where } \rho_i = (1 + s)^{-i}. \tag{A6}$$

Next we calculate x_i , which is the probability that no adaptive variant either already exists at the time of environmental change or appears before drift is complete, given a population of i m_2 individuals at the time of environmental change. We have assumed that adaptive variants that appear before the environmental change event never reproduce, and so the only way an adaptive variant will preexist is if the adaptation appeared at the time of that individual's birth. Then the probability that no adaptive variant preexists is equal to

$$(1 - m_2)^i(1 - m_1)^{N-i}.$$

Let the probability that an adaptive variant appears before fixation by drift is complete be $p(i)$. Then

$$x_i = (1 - m_2)^i(1 - m_1)^{N-i}(1 - p(i)).$$

We present both an exact solution and a more numerically tractable approximate solution for $p(i)$.

Exact solution for $p(i)$: Consider a starting population containing i m_2 individuals. We calculate the probability $p(i)$ that an adaptive variant appears before fixation by drift is complete. This gives us the set of equations

$$\begin{aligned} p(0) &= 0 \\ p(i) &= \frac{i}{N} \left(m_2 + (1 - m_2) \left(\frac{i}{N} p(i) + \frac{N-i}{N} p(i+1) \right) \right) + \frac{N-i}{N} \left(m_1 + (1 - m_1) \left(\frac{i}{N} p(i-1) + \frac{N-i}{N} p(i) \right) \right), \\ i &= 1, \dots, N-1 \\ p(N) &= 0. \end{aligned} \tag{A7}$$

Using the approximate solution described below as an initial estimate, a numerical solution to these equations was found through iteration for small $N = 100$. It was noted that $\log(1 - p(yN))/N$ is reasonably invariant over N for $0 < y < 1$. This invariance was exploited by using the solution for $N = 100$ to produce an initial estimate for a larger value of N . The solution for the larger N was then found by iteration. This process was repeated for successively larger values of N .

Approximate solution for $p(i)$: Let $y = i/N$ and $\phi(i/N) = p(i)$, mapping the discrete function onto a continuous function (IWASA *et al.* 2004). Calculating the Taylor series of $\phi(y)$, we have

$$p(i+1) = \phi\left(y + \frac{1}{N}\right) = \phi(y) + \frac{1}{N} \frac{d\phi}{dy}(y) + \frac{1}{2N^2} \frac{d^2\phi}{dy^2}(y) + \dots \tag{A8a}$$

$$p(i-1) = \phi\left(y - \frac{1}{N}\right) = \phi(y) - \frac{1}{N} \frac{d\phi}{dy}(y) + \frac{1}{2N^2} \frac{d^2\phi}{dy^2}(y) + \dots \tag{A8b}$$

For large N , we can ignore higher-order terms. Note that this may cause problems if $m_1, m_2 = O(1/N)$ is not satisfied, depending on the properties of ϕ . This approximation is motivated by its numerical verification rather than by its rigor. Substituting Equation A8 into Equation A7, we get

$$(\phi(y) - 1)(ym_2 + (1 - y)m_1) = y(1 - y) \left((m_1 - m_2) \frac{1}{N} \frac{d\phi}{dy}(y) + (2 - m_1 - m_2) \frac{1}{2N^2} \frac{d^2\phi}{dy^2}(y) \right). \tag{A9}$$

Assume a solution of the form

$$\phi(y) - 1 = -e^{f(y)N}. \tag{A10a}$$

We then have

$$\phi'(y) = -e^{f(y)N} f'(y) N \tag{A10b}$$

$$\phi''(y) = -e^{f(y)N} f''(y) N - e^{f(y)N} (f'(y))^2 N^2. \tag{A10c}$$

Substituting Equation A10 into Equation A9, we get

$$\frac{ym_2 + (1 - y)m_1}{y(1 - y)} = (m_1 - m_2) f'(y) + (2 - m_1 - m_2) \frac{1}{2} (f''(y)/N + (f'(y))^2).$$

We now have reduced the problem to a first-order differential equation, but unfortunately it still is not easy to solve. Assume that the term $(f'(y))^2$ can be ignored and approximate it as zero. This approximation was motivated solely

by numerical results for the exact solution, and the case for the approximation rests entirely on its numerical verification in Figures 5 and 6. Now integrate this with respect to y to get

$$-m_2 \log(1 - y) + m_1 \log y = (m_1 - m_2) f(y) + \frac{(2 - m_1 - m_2) f'(y)}{2N} + C1. \quad (\text{A11})$$

Equation A11 was solved using Mathematica 4.1 with boundary conditions $f(0) = f(1) = 0$ to give

$$f(y) = \frac{1}{m_1 - m_2} \left(\frac{1}{e^X - 1} \left(-(\gamma + \log|X|) (m_1(1 - e^{X(1-y)}) + m_2 e^X(1 - e^{-Xy})) \right) \right. \\ \left. + m_2 e^{X(1-y)} Ei(-X(1 - y)) - m_1 e^{-Xy} Ei(Xy) - m_2 \log(1 - y) + m_1 \log(y) \right),$$

where $X = 2(m_1 - m_2)N / (2 - m_1 - m_2)$, γ is Euler's constant (≈ 0.577216), and $Ei(x)$ is the exponential integral function $\int_{-x}^{\infty} (e^{-t}/t) dt$. The exponential integral function can readily be evaluated using series expansions.

Numerical note: Note that the summation in Equation 1 is very computationally expensive for large N . In this case, an interpolating approximation was used, with an adaptive algorithm choosing the location and number of points to be evaluated.

